MARKED-UP VERSION OF SUBSTITUTE SPECIFICATION

DEVICE AND METHOD TO COMPRESS DESTINATION ADDRESSES OF A MULTICAST MESSAGE

Background of the Invention

[0001]   The present invention relates to a device for compressing a list of destination addresses of a multicast message, and a method for compressing a list of destination addresses of a multicast message, a router of a communications network wherein such a device for compressing is included, and a host of a communications network wherein such a device for compressing is included.

[0002]   Multicasting is the transmission of one and the same message to a set of destination hosts. Multicasting on the Internet allows to send IP (Internet Protocol) datagrams to every host who cares to listen without the necessity to send the IP datagrams individually to each destination. Broadcasting radio stations on the Internet, audio- or video-conferencing, publishing a newsletter, attending meetings without leaving the office, seeing lectures from professors, and many other applications become realistic without excessive bandwidth occupancy through multicasting. In a host-group implementation of IP multicasting, each multicasted IP packet carries an IP multicast address, which is a single IP destination address that identifies all group members of the multicast session. The membership of the host group is dynamic which means that destination hosts may join and leave the group at any time. The forwarding of IP multicast datagrams is handled by so called multicast routers which are IP routers that have the responsibility to forward IP multicast datagrams to all networks that have members of the destination host group. An overview of multicast protocols for a host implementation of IP multicasting, such as XTP (Xpress Transfer Protocol), ST-II (Experimental Internet Stream Protocol Version 2), MTP (Multicast Transport Protocol), ..., is given in rfc1458 entitled

'Requirements for Multicast Protocols'. This IETF (Internet Engineering Task Force) submission is published on the Internet and can be found at the URL 'http://www.cis.ohio-state.edu/htbin/rfc/rfc1458.html'. The scalability of host implementation of IP multicasting for many groups however is a problem. In particular, the IP backbone network may suffer from scalability problems since the number of groups whose states have to be maintained there can be huge in case a separate multicast address has to be allocated to every small group of hosts attending for example an audio-conference or a video-conference. A connectionless implementation of IP multicasting, such as described in the Internet Draft with reference <draft-ooms-cl-multicast-00.txt> entitled 'Connectionless Multicast', does not have the scalability problem of host implementations of IP multicasting since no state per multicast group is created in the IP routers. The latter Internet Draft, that can be retrieved from the Internet via the URL 'http://info.internet.isi.edu:80/in-drafts/files/draft-ooms-cl-multicast-00.txt', proposes a multicast mechanism wherein each IP packet carries a list of all IP addresses of the multicast session members. In each router, the next hop for each destination in the list is determined by consulting the routing tables, and for every distinct next hop, a new IP header is constructed. This new IP header only contains a list of destinations for which that next hop is on the shortest path to the destinations. The main advantages of connectionless multicasting are that no multicast address allocation is required and routers do not have to maintain a state per session. Connectionless multicasting however requires each packet to contain all remaining destinations. Since the list of destination addresses might be long, connectionless multicasting may significantly increase the overhead and consequently reduce the effective throughput through the Internet.

[0003]   The usefulness of a method and equipment for compressing the list of destination addresses in connectionless multicasted IP packets is expressed in paragraph 6 of the Internet

2

Draft ~~draft-ooms-cl-multicast-00.txt>~~ but a preferred implementation of the compression

technique to be applied is not described therein. The authors of the Internet Draft already

suggested that the compression method preferably should be based on detection of a common

prefix. A common prefix based compression method for IP destination addresses is known for

example from the article 'Fast Address Lookup for Internet Routers' from the authors Stefan

Nilsson and Gunnar Karlsson, ~~retrievable fom the Internet at URL~~

<del>'http://www.nada.kth.se/~nilsson/public/papers/router/'</del>. Therein, a method for compressing the

information held by the routing table of IP routers is described. This compression technique

replaces IP addresses with a prefix and skip value, the skip value representing the number of bits

that can be skipped during a search operation. If in the list of destination addresses of a multicast

message, destination addresses having a common prefix would be replaced with a compound

destination address consisting of the common prefix and a skip number indicative for the number

of irrelevant bits, every host whose address starts with the common prefix would receive the

multicasted message, whether this host forms part of the original list of destinations or not. Thus,

only sets of destination addresses having a common prefix and containing all possible suffixes

could be compressed this way for multicast purposes. This known compression technique based

on common prefixes would therefore be applicable very sparsely and, as a consequence, cannot

significantly reduce the aggregate overhead generated for multicast traffic in the Internet.

Summary of the Invention

[0004]   An aspect ~~object~~ of the present invention is to provide a device and method for

compressing a list of destination addresses that is optimal for multicast purposes in a sense that it

significantly reduces the overhead of multicasted packets in any connectionless multicast session

wherein at least two destination addresses have a common prefix.

[0005]   According to the invention, this aspect object is achieved by the device for compressing a list of destination addresses of a multicast message, the method for compressing a list of destination addresses of a multicast message, the router of a communications network, and the host of a communications network.

[0006]   Indeed, since the suffixes of the destination addresses are explicitly listed in the compound address, any set of destination addresses that has a common prefix can be compressed according to the present invention without any risk that non-members of the multicast session whose address also starts with the common prefix would receive the multicasted message. Moreover, as will be explained later on, compound addresses can further be combined with other destination addresses or other compound addresses, provided they all have a common prefix, to be further compressed into a new compound address. The present compression method in other words can be applied iteratively as long as there exists at least two addresses, destination addresses or compound destination addresses[[,]] that have a common prefix. At each iteration, the overhead of the multicasted packets is further reduced.

[0007]   It is to be noticed that the term 'comprising', used in the claims, should not be interpreted as being limitative to the means listed thereafter. Thus, the scope of the expression 'a device comprising means A and B' should not be limited to devices consisting only of components A and B. It means that with respect to the present invention, the only relevant components of the device are A and B.

[0008]   Thus, any subset of IP addresses that has a common prefix can be compressed in accordance with the present invention.

[0009]   Thus, if the list of destination addresses contains compound addresses and IP addresses that have a common prefix, a new compound address can be generated wherein the

4

compound addresses and IP addresses are compressed. The current invention consequently can be applied iteratively so that first, a list of IP destination addresses with a rather large common prefix is compressed into a compound address. This compound address can then further be combined with other compound addresses or IP destination addresses that all have a shorter common prefix to generate a new compound address in a next compression iteration. Iterative application of the compression technique according to the present invention will be illustrated by means of a concrete example later on in this patent application.

[0010] Thus, a few compound addresses generated at a first compression iteration, may still have a common prefix and consequently may further be compressed into a new compound address at a second compression iteration.

[0011] In this way, the list of destination addresses of multicasted packets may be compressed before transmission thereof, thus reducing the overhead for connectionless multicasting on the outgoing links of the sourcing host.

[0012] In this way, the router that constructs a new header for each distinct next hop, can compress the list of destinations for which that next hop is on the shortest path to these destinations also in accordance with the present invention. In this way, the overhead transferred over the outgoing links of routers that forward connectionless multicasted packets is also significantly reduced.

[0013] Thus, the compound addresses not only reduce the overhead required to transfer connectionless multicasted messages to the different destinations, but also enable fast access to the routing tables in the routers. This is so because the routing tables do not have to be addressed separately for destinations whose addresses have a common prefix. If the table lookup mechanism involves testing address bits ordered from left to right (where left corresponds to the

5

most significant part, i.e. the prefix), the common prefix of the compound address should be looked up only once as a result of which the routing table lookup performance increases.

Brief Description of the Drawing

[0014]   The above mentioned and other aspects ~~objects~~ and features of the invention will become more apparent and the invention itself will be best understood by referring to the following description of an embodiment taken in conjunction with the accompanying drawing Fig. 1 which is a scheme of an Internet INTERNET wherein the method for compressing a list of destination addresses of a multicast message according to the present invention is implemented.

Detailed Description of the Invention

[0015]   In Fig. 1, four hosts, H1, D1, D2 and D3, and three routers, R1, R2 and R3, of the Internet INTERNET are drawn. Host H1 is connected to a port of the first router R1 via link L11. The ports R1P1 and R1P2 of the first router R1 are interconnected with ports of respectively the second router R2 and the third router R3 via respectively the link L12 and the link L13. Link L21 connects port R2P1 of the second router R2 to a port of the host D1. Similarly, link L22 connects port R2P2 of the second router R2 to a port of host D2 and link L33 interconnects port R3P1 of router R3 with a port of host D3. As is indicated on Fig. 1, host D1 has address A.B.C.D, host D2 has address A.B.C.E and host D3 has address A.F.G.H. In these addresses, each letter is supposed to represent an octet so that each address consists of 32 bits (4 octets). Host H1 will play the role of sourcing host in the example described below so that its address does not need to be known in order to be able to illustrate the compression method according to

6

the present invention. Host H1 and the three routers R1, R2 and R3 are supposed to incorporate a destination list compression device according to the present invention.

[0016]   To explain the invented compression technique it is supposed that host H1 has to multicast an IP (Internet Protocol) datagram to the destination hosts D1, D2 and D3 and thereto applies connectionless multicasting. In the overhead section of this IP datagram, host H1 thus has to identify the destination hosts D1, D2 and D3 by their respective IP addresses A.B.C.D, A.B.C.E and A.F.G.H. The destination list compression device in host H1 will aid to realize this with low overhead consumption. The destination list compression device in host H1 detects that the addresses A.B.C.D and A.B.C.E of respectively host D1 and host D2 hove a common prefix A.B.C. By subtracting this common prefix A.B.C from the addresses A.B.C.D and A.B.C.E, the compression device obtains the suffixes D and E which it uses to generate a suffix list {D,E}. This suffix list {D,E} is added to the common prefix A.B.C to constitute a compound address A.B.C{D,E} that still indicates that the two hosts D1 and D2 belong to the destinations of the IP datagram but which contains only 5 octets, i.e., A, B, C, D and E, instead of the 8 octets, A, B, C, D, A, B, C and E, that have to be embedded in the IP datagram overhead if no compression is applied. As a result of the first iteration step in the compression method, host H1 obtains a list of destination addresses for the IP datagram to be multicasted that consists of the IP address A.F.G.H and the compound destination address A.B.C{D,E}. In a second iteration step, the compression device in host H1 detects that the IP address A.F.G.H and the compound address A.B.C{D,E} still have a common prefix A. By subtracting this common prefix A from the IP address A.F.G.H and the compound address A.B.C{D,E}, the compression device of host H1 generates the suffixes F.G.H and B.C{D,E} from which the list of suffixes {B.C{D,E},F.G.H} is constituted. This list of suffixes {B.C{D,E},F.G.H} is added to the common prefix A to

generate a new compound address A{B.C{D,E},F.G.H} that indicates that the IP datagram has to be multicasted to the destination hosts D1, D2 and D3, but which thereto occupies only 8 octets, i.e., A, B, C, D, E, F, G, H, instead of the 12 octets, A, B, C, D, A, B, C, E, A, F, G and H, that would have been embedded in the overhead section of the IP datagram if no compression was applied. In this way, the overhead for transferring the IP datagram over link L11 has been reduced significantly.

[0017]   Router R1, upon receipt of the IP datagram addresses its routing table with the compound address A{B.C{D,E},F.G.H}. Such a routing table lookup involves testing the address bits ordered from left to right and is shortened if the compound address A{B.C{D,E},F.G.H} is used instead of the three addresses A.B.C.D, A.B.C.E and A.F.G.H because the common prefixes of the addresses, A and A.B.C, have to be looked up only once via the compound address A{B.C{D,E},F.G.H} instead of respectively three or two times. Routing table lookup performance hence is increased. From the routing table in router R1 it is derived that the destinations whose addresses start with prefix A.B.C, e.g. A.B.C.D and A.B.C.E, have router R2 as next hop on the shortest path thereto. The destination with address A.F.G.H has router R3 as next hop on the shortest path thereto. Router R1 consequently will constitute a new IP datagram to be forwarded via link L12 to router R2 and a new IP datagram to be forwarded via link L13 to router R3. The payload sections of these new IP datagrams are copies of the payload section of the IP datagram received via link L11. In the overhead section of the IP datagram that will be forwarded over link L12, the list of addresses A.B.C.D and A.B.C.E has to be included. The compressed version of this list, A.B.C{D,E}, will be copied from the IP datagram received via link L11 into the header of the new IP datagram that will be transferred over link L12. Alternatively, this list will be re-compressed by the compression device in router

8

R1. This compression device then detects the common prefix A.B.C in the addresses A.B.C.D and A.B.C.E. By subtracting this common prefix from the addresses A.B.C.D and A.B.C.E, the suffixes D and E are obtained, which are combined into a list of suffixes {D,E}. The latter list of suffixes {D,E} is added to the common prefix A.B.C to generate the compound address A.B.C{D,E} that will be embedded in the IP datagram that is forwarded via link L12 to router R2. In the overhead section of the IP datagram transferred over L12, the number of octets required to indicate that this IP datagram has to be forwarded to the destination hosts D1 and D2 is reduced from 8 to 5. In the header of the IP datagram that will be forwarded by router R1 over L13 to router R3, router R1 embeds the address A.F.G.H of destination hosts D3 which cannot be compressed anymore.

[0018]   Router R2, upon receipt of the IP datagram forwarded over the link L12, addresses its routing table with the compound address A.B.C{D,E} subtracted from the header of this IP datagram. From the routing table in router R2 it is derived that the address A.B.C.D can be reached via port R2P1 and that the address A.B.C.E can be reached via port R2P2. Router R2 consequently will constitute a new IP datagram in whose header the address A.B.C.D will be embedded, and another new IP datagram in whose header the address A.B.C.E will be embedded. The payload sections of these new IP datagrams are again copies. The first IP datagram constituted by router R2 is forwarded via port R2P1 and link L21 to the destination host D1 whereas the second IP datagram constituted by router R2 is forwarded via port R2P2 and link L22 to the destination host D2. The addresses A.B.C.D and A.B.C.E cannot be compressed anymore although router R2 is supposed to incorporate a compression device according to the present invention.

9

[0019]   Router R3, upon receipt of the IP datagram transferred over the link L13, addresses its routing table with the address A.F.G.H and derives from its routing table that the received IP datagram has to be forwarded via port R3P1 and link L33 to the destination host D3. Since there is only one next hop, this IP datagram can simply be copied and forwarded by router R3. No compression of the destination address is possible, although it is supposed that also router R3 is equipped with a compression device according to the present invention.

[0020]   Summarizing, iterative application of the compression technique according to the present invention, allowed to reduce the overhead section that is indicative for the destination addresses of the multicasted IP datagram from 12 to 8 octets on link L11. On link L12 between router R1 and router R2, this overhead section could be reduced from 8 to 5 octets. It is clear that statistically the obtained reduction of overhead will be even more significant in case more destination hosts are member of the multicast session, and in case these destinations have longer common prefixes.

[0021]   Although the above described embodiment concerns routing of IP (Internet Protocol) datagrams in the Internet, it will be obvious to any person skilled in the art of communication networks, that the application of the compression technique for a multicast destination address list is not limited to any particular network or internetwork, and accordingly also not to any particular format of the data packets. This invention in other words is more widely applicable than in the Internet.

[0022]   It is also noticed that the compression technique according to the present invention does not necessarily have to be applied octet-aligned as described in the above example. The compression technique alternatively for example can be applied nibble-aligned or bit-aligned.

[0023]   Furthermore, it is remarked that an embodiment of the present invention is described above rather in terms of functional blocks. From the functional description of these blocks it will be obvious for a person skilled in the art of designing electronic devices how embodiments of these blocks can be manufactured with well-known electronic components. A detailed architecture of the contents of the functional blocks hence is not given.

[0024]   While the principles of the invention have been described above in connection with specific apparatus, it is to be clearly understood that this description is made only by way of example and not as a limitation on the scope of the invention.